A new asymptotic enumeration technique: the Lovász Local Lemma

Linyuan Lu [*]   László Székely [†]

University of South Carolina {lu, szekely}@math.sc.edu

First Draft

May 25, 2009

Abstract

Our previous paper [14] applied a general version of the Lovász Local Lemma that allows negative dependency graphs [11] to the space of random injections from an m-element set to an n-element set. Equivalently, the same story can be told about the space of random matchings in $K_{n,m}$. Now we show how the cited version of the Lovász Local Lemma applies to the space of random matchings in K_{2n}. We also prove tight upper bounds that asymptotically match the lower bound given by the Lovász Local Lemma. As a consequence, we give new proofs to results on the enumeration of d-regular graphs. The tight upper bounds can be modified to the space of matchings in $K_{n,m}$, where they yield as application asymptotic formulas for permutation and Latin rectangle enumeration problems.

As another application, we provide a new proof to the classical probabilistic result of Erdős [8] that showed the existence of graphs with arbitrary large girth and chromatic number. In addition to letting the girth and chromatic number slowly grow to infinity in terms of the number of vertices, we provide such a graph with a prescribed degree sequence, if the degree sequence satisfies some mild conditions.

1   Lovász Local Lemma with negative dependency graphs

This is a sequel to our previous paper [14] and we use the same notations. Let $A_1, A_2, \ldots, A_n$ be events in a probability space.

A *negative dependency graph* for $A_1, \ldots, A_n$ is a simple graph on $[n]$ satisfying

$$\Pr(A_i \mid \wedge_{j \in S} \overline{A_j}) \leq \Pr(A_i), \tag{1}$$

---

for any index i and any subset $S \subseteq \{j \mid ij \notin E(G)\}$, if the conditional probability $\Pr(A_i \mid \wedge_{j \in S} \overline{A_j})$ is well-defined, i.e. $\Pr(\wedge_{j \in S} \overline{A_j}) > 0$. We will make use of the fact that inequality (1) trivially holds when $\Pr(A_i) = 0$, otherwise the following inequality is equivalent to inequality (1):

$$\Pr(\wedge_{j \in S} \overline{A_j} \mid A_i) \leq \Pr(\wedge_{j \in S} \overline{A_j}). \tag{2}$$

For variants of the Lovász Local Lemma with increasing strength, see [10, 18, 11, 13]:

Lemma 1 [Lovász Local Lemma.] *Let* $A_1, \ldots, A_n$ *be events with a negative dependency graph* G. *If there exist numbers* $x_1, \ldots, x_n \in [0, 1)$ *such that*

$$\Pr(A_i) \leq x_i \prod_{ij \in E(G)} (1 - x_j) \tag{3}$$

for all i, *then*

$$\Pr(\wedge_{i=1}^n \overline{A_i}) \geq \prod_{i=1}^n (1 - x_i). \tag{4}$$

The main obstacle for using Lemma 1 is the difficulty to define a useful negative dependency graph other than a dependency graph. In [14], we described a general way to create negative dependency graphs in the space of random functions $U \to V$ equipped with uniform distribution. Namely, let the events be the set of all extensions of some particular partial functions to functions; and create an edge for the negative dependency graph, if the partial functions have common elements in their domains or ranges, other than the agreement of the partial functions. These events also can be thought of as all extensions of (partial) matchings in the complete bipartite graph with classes U, V, where an edge of the negative dependency graph comes from two event-defining (partial) matchings whose union is no longer a (partial) matching after suppressing multiple edges. In [14], we used this technique to prove a new result on the Turán hypergraph problem, and we found surprising applications as proving lower bounds (matching certain asymptotic formulas) for permutation and Latin rectangle enumeration problems.

In this paper, we show an analogous construction of a negative dependency graph for events, which live in the space of random matchings of a complete graph. We require that the events are the set of all extensions of (partial) matchings in a complete graph to perfect matchings, and two event-defining partial matchings make an edge, if their union is no longer a (partial) matching after suppressing multiple edges. (Although our construction fails for extensions of partial matchings of arbitrary graphs, there might be some other graph classes providing interesting results.)

We move one step further and show some general and some specific *upper bounds* for the event estimated by the Lovász Local Lemma, and show that for large classes of problems the upper bound is asymptotically equal to the lower bound. These results apply to the permutation enumeration problems in [14],

and to enumeration problems for regular graphs. The asymptotic enumeration results that we prove are not new—with one exception—and typically do not give the largest known range of the asymptotic formula, but are nontrivial results and often more recent than the Lovász Local lemma itself. They come out from our framework elementarily, and the lower bounds even easily.

In a forthcoming paper we will extend our negative dependency graph construction to matchings in complete r-uniform hypergraphs and will apply this result to hypergraph enumeration.

As another application, we revisit a classic of the probabilistic method: Erdős' proof to the existence of graphs with arbitrary large girth and chromatic number [8]. We show that in a wide range the degree sequence of this graph can be prescribed as well, without requiring much bigger graphs.

In a scenario of the Poisson paradigm, we estimate the probability that none of a set of rare events occur. Let X be the sum of the indicator variables of these events and $\mu = \mathrm{E}(X)$. If the dependency among these events is rare, then one would expect that X has a Poisson distribution with mean μ. In particular, $\Pr(X = 0) \approx e^{-\mu}$. The Janson inequality and Brun's sieve method [1] are often the good choice to solve these kind of problems. Now we offer an alternative approach—using Lovász Local Lemma. Our approach can be considered as an analogue of the Janson inequality in another setting that offers plenty of applications. It is curious that the proof of Boppana and Spencer [5] for the Janson inequality (see also in [1]) uses conditional probabilities somewhat similarly to the proof of the Lovász Local Lemma.

For further research, it would be interesting to get asymptotically further terms from the Poisson distribution for the probability of exactly k events holding, for any fixed k. It would also be interesting to get a few more terms from the asymptotic expansion of $Pr(\wedge_{i=1}^{n} \overline{A_i})$ in the matching models. Also, a number of further applications are possible.

2 Some general results on negative and near-positive dependency graphs

These lower and upper bounds are *general* in the sense that there is no assumption on the events being defined through matchings.

In many applications, where $n \to \infty$, both $\Pr(A_i)$ and $\sum_{ij \in E(G)} \Pr(A_j)$ are so small that one can set $x_i =: (1 + o(1))\Pr(A_i)$ to use Lemma 1. We are going to show this in some generality, to give a taste of the later results. More precise lower bound in a specific setting will be given in Theorem 4.

Lemma 2 *Let $A_1, \ldots, A_n$ be events with negative dependency graph G. Let us be given any ϵ with $0 < \epsilon < 0.14$. If $\Pr(A_i) < \epsilon$ and $\sum_{ij \in E(G)} \Pr(A_j) < \epsilon$ for every $1 \le i \le n$, then*

$$\Pr(\wedge_{i=1}^{n} \overline{A_i}) \ge e^{-(1+3\epsilon) \sum_{i=1}^{n} \Pr(A_i)}.$$

Proof: Set $x_i = (1+3\epsilon)\mathrm{Pr}(A_i)[1-\mathrm{Pr}(A_i)]$. It is clear that $0 \leq x_i < 1$. Observe that for every i

$$
\begin{aligned}
1 - x_i &= 1 - (1+3\epsilon)[\mathrm{Pr}(A_i) - \mathrm{Pr}(A_i)^2] \\
&\geq 1 - (1+3\epsilon)\mathrm{Pr}(A_i) + \frac{[(1+3\epsilon)\mathrm{Pr}(A_i)]^2}{2!} \quad (5)
\end{aligned}
$$

and by the Taylor expansion of the exponential function

$$
\geq e^{-(1+3\epsilon)\mathrm{Pr}(A_i)}. \quad (6)
$$

Based on (5, 6), to prove condition (2), we have to show only the first inequality in

$$\mathrm{Pr}(A_i) \leq x_i e^{-(1+3\epsilon)\epsilon} \leq x_i \prod_{ij \in E(G)} e^{-(1+3\epsilon)\mathrm{Pr}(A_j)} \leq x_i \prod_{ij \in E(G)} (1 - x_j).$$

The first inequality follows from the the upper bound on ϵ, as $1 < (1-\epsilon)(1+3\epsilon)e^{-(1+3\epsilon)\epsilon}$ in the interval $(0, 0.14)$. Finally, the conclusion follows from multiplying out the inequalities (5, 6) for all i. $\square$

Next we give a crucial new definition. For the events $A_1, \ldots, A_n$ in a probability space Ω, and an ϵ with $1 > \epsilon > 0$, we define an ϵ-*near-positive dependency graph* to be a graph G on $V(G) = [n]$ satisfying

1. $\mathrm{Pr}(A_i \wedge A_j) = 0$ if $ij \in E(G)$.

2. For any index i and any subset $T \subseteq \{j \mid ij \notin E(G)\}$,

$$\mathrm{Pr}(A_i \mid \wedge_{j \in T} \overline{A_j}) \geq (1-\epsilon)\mathrm{Pr}(A_i),$$

   whenever the conditional probability is well-defined.

Theorem 1 *Let $A_1, \ldots, A_n$ be events with an ϵ-near-positive dependency graph G. Then we have*

$$\mathrm{Pr}(\wedge_{i=1}^n \overline{A_i}) \leq \prod_{i=1}^n [1 - (1-\epsilon)\mathrm{Pr}(A_i)].$$

Proof: If $\mathrm{Pr}(\wedge_{i=1}^n \overline{A_i}) = 0$, then the conclusion holds. So we may assume without loss of generality that $\mathrm{Pr}(\wedge_{i=1}^n \overline{A_i}) > 0$. Now we would like to show that for any i and any subset $S \subseteq V(G)$ with $i \notin S$,

$$\mathrm{Pr}(A_i \mid \wedge_{j \in S} \overline{A_j}) \geq (1-\epsilon)\mathrm{Pr}(A_i),$$

as the conditional probability above is well-defined by our assumption. Write $S = S_1 \cup S_2$, where $S_1 = S \cap N_G(i)$ and $S_2 = S \setminus S_1$. We have

$$
\begin{aligned}
\mathrm{Pr}(A_i \mid \wedge_{j \in S} \overline{A_j}) &= \frac{\mathrm{Pr}(A_i \wedge (\wedge_{k \in S_1} \overline{A_k}) \mid \wedge_{j \in S_2} \overline{A_j})}{\mathrm{Pr}(\wedge_{k \in S_1} \overline{A_k} \mid \wedge_{j \in S_2} \overline{A_j})} \\
&= \frac{\mathrm{Pr}(A_i \mid \wedge_{j \in S_2} \overline{A_j})}{\mathrm{Pr}(\wedge_{k \in S_1} \overline{A_k} \mid \wedge_{j \in S_2} \overline{A_j})} \\
&\geq \mathrm{Pr}(A_i \mid \wedge_{j \in S_2} \overline{A_j}) \\
&\geq (1-\epsilon)\mathrm{Pr}(A_i).
\end{aligned}
$$

(The first part of the definition of the ϵ-near-positive dependency graph, $\Pr(A_i \wedge A_j) = 0$ for ij edges, allowed the elimination of the $\wedge_{k \in S_1} \overline{A_k}$ term.) Hence, we have

$$
\begin{aligned}
\Pr(\wedge_{i=1}^n \overline{A_i}) &= \prod_{i=1}^n \Pr(\overline{A_i} \mid \wedge_{k=i+1}^n \overline{A_k}) = \\
\prod_{i=1}^n [1 - \Pr(A_i \mid \wedge_{k=i+1}^n \overline{A_k})] &\leq \prod_{i=1}^n (1 - (1-\epsilon)\Pr(A_i)).
\end{aligned}
$$

$\square$

Lemma 3 *Assume that G is a negative dependency graph for the events $A_1, A_2, ..., A_n$. Assume further that $V(G)$ has a partition into classes, such that any two events in the same class have empty intersection. For any partition class J, let $B_J = \vee_{j \in J} A_j$. Now the quotient graph of G is a negative dependency graph for the events B_J. Furthermore, if G is an ϵ-near positive dependency graph, and the quotient graph has no edges, then the quotient graph is also an ϵ-near positive dependency graph.*

Proof. We proved the first part of the Lemma in [14], for completeness we include the proof. We have to show that if $\mathcal{K}$ is a subset of non-neighbors of J in the quotient graph, then $\Pr(B_J \mid \wedge_{K \in \mathcal{K}} \overline{B_K}) \leq \Pr(B_J)$. By the additivity of (conditional) probability over mutually exclusive events, and it is sufficent to show that

$$
\Pr(A_j \mid \wedge_{K \in \mathcal{K}} \overline{B_K}) \leq \Pr(A_j), \tag{7}
$$

and by formula (2) equivalently:

$$
\Pr(\wedge_{K \in \mathcal{K}} \overline{B_K} \mid A_j) \leq \Pr(\wedge_{K \in \mathcal{K}} \overline{B_K}) \tag{8}
$$

holds for every $j \in J$. However, $\wedge_{K \in \mathcal{K}} \overline{B_K} = \wedge_{i \in \cup \mathcal{K}} \overline{A_i}$, and every $i \in \cup \mathcal{K}$ is a non-neighbor of j in G, according to the definition of the quotient graph. Therefore, (8) holds as G is a negative dependency graph. To prove the second part, note that 1. in the first condition to be an ϵ-near positive dependency graph holds, as there are no edges in the quotient graph. To prove 2., one has to show that $\Pr(B_J \mid \wedge_{K \in \mathcal{K}} \overline{B_K}) \geq (1-\epsilon)\Pr(B_J)$ for any $\mathcal{K}$ and J, with $J \notin \mathcal{K}$. An argument like in the first part works. $\square$

3 Examples for negative dependency graphs: The space of random matchings of K_N and $K_{N,M}$

For an even integer N, Let Ω denote the probability space of perfect matchings of the complete graph K_N for an even integer N; or the probability space of perfect matchings of the complete bipartite graph $K_{N,M}$ equipped with the uniform distribution. We are going to apply the Lovász Local Lemma (Lemma

1) in Ω by identifying a class of negative dependency graphs. For any (not necessary perfect) matching M, let A_M be the set of perfect matching extending M:

$$A_M = \{F \in \Omega \mid M \subseteq F\}. \tag{9}$$

We will term an event A_M in (9), with $M \neq \emptyset$, a *canonical event.* We will say that two matchings, M_1 and M_2, are in *conflict,* if $M_1 \cup M_2$ is not a matching after suppressing multiple edges.

Theorem 2 *Let $\mathcal{M}$ be a collection of matchings in K_N or $K_{N,M}$. The graph G described below is a negative dependency graph for the canonical events $\{A_M \mid M \in \mathcal{M}\}$:*

- $V(G) = \mathcal{M}$,

- $E(G) = \{M_1 M_2 \mid M_1 \in \mathcal{M} \text{ and } M_2 \in \mathcal{M} \text{ are in conflict}\}$.

Proof: For complete bipartite graphs we proved this theorem in [14], and therefore we have to prove it now for K_N. We will prove the theorem by induction on N. The base case $N = 2$ is trivial. Throughout this paper, we always assume that the vertex set of K_N is $[N] = \{1, 2, \ldots, N\}$. There is a canonical injection from $[N]$ to $[N + s]$, and consequently from $V(K_N)$ to $V(K_{N+s})$ and from $E(K_N)$ to $E(K_{N+s})$. Through this canonical injection, every matching of K_N can be viewed as a matching of K_{N+s}. (Note that a perfect matching in K_N will not be perfect in K_{N+s} for $s > 0$.) To emphasize the difference in the vertex set, we use A_M^N to denote the event of $\Omega = \Omega_N$ induced by the matching M.

Lemma 4 *For any collection $\mathcal{M}$ of matchings in K_N, we have*

$$\Pr(\wedge_{M \in \mathcal{M}} \overline{A_M^N}) \leq \Pr(\wedge_{M \in \mathcal{M}} \overline{A_M^{N+2}}).$$

Proof: We partition the space of Ω_{N+2} into $N + 1$ sets as follows: for $1 \leq i \leq N + 1$, let $\mathcal{C}_i$ be the set of perfect matchings containing the edge $i(N + 2)$. We have

$$\Pr(\wedge_{M \in \mathcal{M}} \overline{A_M^{N+2}}) = \sum_{i=1}^{N+1} \Pr(\wedge_{M \in \mathcal{M}} \overline{A_M^{N+2}} \wedge \mathcal{C}_i).$$

We observe that $\mathcal{C}_i \subseteq \overline{A_M^{N+2}}$ if and only if M conflicts $i(N + 2)$, a one-edge matching. Let $\mathcal{B}_i$ be the subset of $\mathcal{M}$, whose elements are not in conflict with the edge $i(N + 2)$. (In particular, $\mathcal{B}_{N+1} = \mathcal{M}$.) We have

$$\wedge_{M \in \mathcal{M}} \overline{A_M^{N+2}} \wedge \mathcal{C}_i = \wedge_{M \in \mathcal{B}_i} \overline{A_M^{N+2}} \wedge \mathcal{C}_i.$$

Let ϕ_i be the transposition $i \leftrightarrow N+1$. Note that ϕ_i stabilizes $\mathcal{B}_i$, interchanges $\mathcal{C}_i$ and $\mathcal{C}_{N+1}$, and maps $\wedge_{M \in \mathcal{B}_i} \overline{A_M^{N+2}} \wedge \mathcal{C}_i$ to $\wedge_{M \in \mathcal{B}_i} \overline{A_M^{N+2}} \wedge \mathcal{C}_{N+1}$. We have

$$
\begin{aligned}
\Pr(\wedge_{M \in \mathcal{M}} \overline{A_M^{N+2}}) &= \sum_{i=1}^{N+1} \Pr(\wedge_{M \in \mathcal{M}} \overline{A_M^{N+2}} \wedge \mathcal{C}_i) \qquad (10) \\
&= \sum_{i=1}^{N+1} \Pr(\wedge_{M \in \mathcal{B}_i} \overline{A_M^{N+2}} \wedge \mathcal{C}_i) \\
&= \sum_{i=1}^{N+1} \Pr(\wedge_{M \in \mathcal{B}_i} \overline{A_M^{N+2}} \wedge \mathcal{C}_{N+1}) \\
&= \sum_{i=1}^{N+1} \Pr(\wedge_{M \in \mathcal{B}_i} \overline{A_M^{N+2}} \mid \mathcal{C}_{N+1}) \Pr(\mathcal{C}_{N+1}) \\
&= \frac{1}{N+1} \sum_{i=1}^{N+1} \Pr(\wedge_{M \in \mathcal{B}_i} \overline{A_M^{N}}), \qquad (11)
\end{aligned}
$$

and estimating further

$$
\begin{aligned}
&\geq (N+1) \Pr(\wedge_{M \in \mathcal{M}} \overline{A_M^{N}}) \frac{1}{N+1} \\
&= \Pr(\wedge_{M \in \mathcal{M}} \overline{A_M^{N}}).
\end{aligned}
$$

The proof of Lemma 4 is finished. $\qquad \square$

We are back to the proof of Theorem 2: For any fixed matching $M \in \mathcal{M}$, and a subset $\mathcal{J} \subseteq \mathcal{M}$ satisfying that for every $M' \in \mathcal{J}$, M' is not in conflict with M, by (2) it suffices to show that

$$
\Pr(\wedge_{M' \in \mathcal{J}} \overline{A_{M'}} \mid A_M) \leq \Pr(\wedge_{M' \in \mathcal{J}} \overline{A_{M'}}). \qquad (12)
$$

Let $\mathcal{J}' = \{M' \setminus M \mid M' \in \mathcal{J}\}$. Assume first that $\emptyset \notin \mathcal{J}'$. Since every matching M' in $\mathcal{J}$ is not in conflict with M, the vertex set $V(M' \setminus M)$ of $M' \setminus M$ is disjoint from the vertex set $V(M)$ of M. Let $T = V(M)$ be the set of vertices covered by the matching M and U be the set of vertices covered by at least one matching $F \in \mathcal{J}'$. We have $T \cap U = \emptyset$. Let π be the permutation of S_N mapping T to $\{N - |T| + 1, N - |T| + 2, \dots, N\}$. We have $\pi(U) \cap \pi(T) = \emptyset$. Thus, $\pi(U) \subseteq [N - |T|]$. Let $\pi(\mathcal{J}') = \{\pi(F) \mid F \in \mathcal{J}'\}$ and $F' = \pi(F)$. Each matching in $\pi(\mathcal{J}')$ is a matching in $K_{N-|T|}$. We obtain (12) using Lemma 4

repeatedly:

$$
\begin{aligned}
\Pr(\wedge_{M' \in \mathcal{J}} \overline{A_{M'}} \mid A_M) &= \frac{\Pr(\wedge_{M' \in \mathcal{J}} \overline{A_{M'}} \wedge A_M)}{\Pr(A_M)} \\
&= \frac{\Pr(\wedge_{M' \in \mathcal{J}} \overline{A_{M' \setminus M}} \wedge A_M)}{\Pr(A_M)} \\
&= \frac{\Pr(\wedge_{F \in \mathcal{J}'} \overline{A_F} \wedge A_M)}{\Pr(A_M)} \\
&= \Pr(\wedge_{F \in \mathcal{J}'} \overline{A_F} \mid A_M) \\
&= \Pr(\wedge_{F' \in \pi(\mathcal{J}')} \overline{A_{F'}^N} \mid A_{\pi(M)}) \\
&= \Pr(\wedge_{F' \in \pi(\mathcal{J}')} \overline{A_{F'}^{N-|T|}}) \\
&\leq \Pr(\wedge_{F' \in \pi(\mathcal{J}')} \overline{A_{F'}^N}) \\
&= \Pr(\wedge_{F \in \mathcal{J}'} \overline{A_F^N}) \\
&= \Pr(\wedge_{M' \in \mathcal{J}} \overline{A_{M' \setminus M}^N}) \\
&\leq \Pr(\wedge_{M' \in \mathcal{J}} \overline{A_{M'}^N}).
\end{aligned}
$$

If $\emptyset \in \mathcal{J}'$, then the LHS of the estimate above is zero, and therefore we have nothing to do. $\qquad\square$

The following example shows that in Theorem 2 one cannot have an arbitrary graph in the place of K_N or $K_{m,n}$. Consider $G = C_6$, this graph has two perfect matchings. Let e and f denote two opposite edges of C_6. Consider the following two partial matchings: $\{e\}$ and $\{f\}$. We have $Pr(A_{\{e\}}) = Pr(A_{\{f\}}) = 1/2$. However, we have

$$
Pr(A_{\{e\}} \mid \overline{A_{\{f\}}}) = \frac{Pr(A_{\{e\}} \wedge \overline{A_{\{f\}}})}{Pr(\overline{A_{\{f\}}})} \nleq Pr(A_{\{e\}}).
$$

4 Upper bounds in the matching models

Now we consider Ω, the uniform probability space of perfect matchings in K_N (N even) or $K_{N,M}$ (with $M \leq N$). Let $\mathcal{M}$ be a collection of partial matchings. Let $\mathcal{M}_i$ denote the subset of matchings in $\mathcal{M}$ of size i, i.e. for $M \in \mathcal{M}$, we have $M \in \mathcal{M}_i$ if and only if $|M| = i$. Define $I = \{i : \mathcal{M}_i \neq \emptyset\}$. Define $r = \max_{i \in I} i = \max_{M \in \mathcal{M}} |M|$, the maximum size of elements of $\mathcal{M}$. Also, let d_i be an upper bound for the maximum degree of vertices with respect to the edge multiset of $\mathcal{M}_i$, i.e.

$$
d_i \geq \max_{v \in [N]} |\{M \mid v \in V(M), \ M \in \mathcal{M}_i\}|.
$$

We will use the notation $p_{N,i} = \frac{1}{(N-1)(N-3)\cdots(N-2i+1)}$ in K_N and $p_{N,i} = \frac{1}{N(N-1)\cdots(N-i+1)}$ in $K_{N,M}$ We will use the notation introduced here freely in

this section. For any $F \in \mathcal{M}$, let

$$\mathcal{M}_F = \{M \setminus F \mid M \in \mathcal{M}, M \neq F, M \cap F \neq \emptyset, F \text{ is not in conflict to } M\}.$$

We say that a matching $\mathcal{M} = \sum_{i \in I} \mathcal{M}_i$ is δ-*sparse* if

1. No matching from $\mathcal{M}$ is a subset of another matching from $\mathcal{M}$.

2. $\sum_{j \in I} d_j p_{N-2r+2,j} < \frac{1}{8r} - \delta$.

3. For any $F \in \mathcal{M}$, $\sum_{M \in \mathcal{M}_F} p_{N,|M|} \leq \delta$.

4. $16r\delta < 1$.

The main result of this section is the following theorem.

Theorem 3 *Let $\mathcal{M}$ be a collection of matchings in K_N or $K_{N,M}$. If $\mathcal{M}$ is δ-sparse, then the negative dependency graph is also an ϵ-near-positive dependency graph with*

$$\epsilon = 1 - (1 - 2\delta) \prod_{i=0}^{r-1} \frac{1 + \sqrt{1 - 8r \sum_{j \in I} d_j P_{N-2i,j}}}{2}; \qquad (13)$$

and therefore

$$\Pr(\wedge_{M \in \mathcal{M}} \overline{A_M}) \leq \prod_{M \in \mathcal{M}} \left(1 - \Pr(A_M)(1 - 2\delta) \prod_{i=0}^{r-1} \frac{1 + \sqrt{1 - 8r \sum_{j \in I} d_j p_{N-2i,j}}}{2} \right). \tag{14}$$

We are going to prove Theorem 3 for K_N, and leave the proof for $K_{n,m}$, which requires only negligible changes, to the Reader.

Lemma 5 *Let $\mathcal{M}$ be a collection of matchings in K_N. For any $S, \mathcal{T} \subseteq \mathcal{M}$ with $S \cap \mathcal{T} = \emptyset$, we have*

$$\Pr(\wedge_{M \in S} \overline{A_M} \mid \wedge_{M \in \mathcal{T}} \overline{A_M}) \geq \prod_{i \in I} \left(1 - \frac{2p_{N,i}}{1 + \sqrt{1 - 8r \sum_{j \in I} d_j p_{N,j}}} \right)^{|S \cap \mathcal{M}_i|}. \tag{15}$$

Proof: Let G be the negative dependency graph for the family of events $\{A_M\}_{M \in \mathcal{M}}$ according to Theorem 2. For any $M \in \mathcal{M}_i$ and $j \in I$, M conflicts at most $2id_j$ of other matchings in $\mathcal{M}_j$. As for $M \in \mathcal{M}_i$ we have $\Pr(A_M) = p_{N,i}$, in order to apply the Lovász Local Lemma (Lemma 1), we would find $0 \leq x_i < 1$ numbers satisfying

$$p_{N,i} \leq x_i \prod_{j \in I} (1 - x_j)^{2id_j}.$$

9

(Note that $j = i$ is included in the product.) Choose $x_i = yp_{N,i}$ for $y = \frac{2}{1+\sqrt{1-8r\sum_{j\in I} d_j p_{N,j}}}$, the largest root of

$$1 = y(1 - 2ry\sum_{j\in I} d_j p_{N,j}).$$

Note that $y < 2$, and therefore $x_i < 1$ for $N > 2$. For any $i \in I$, we have

$$
\begin{aligned}
p_{N,i} &= \frac{x_i}{y} \\
&= x_i(1 - 2ry\sum_{j\in I} d_j p_{N,j}) \\
&= x_i(1 - 2r\sum_{j\in I} d_j x_j) \\
&\leq x_i \prod_{j\in I}(1 - x_j)^{2rd_j} \text{ (by Bernoulli's inequality)} \\
&\leq x_i \prod_{j\in I}(1 - x_j)^{2id_j}.
\end{aligned}
$$

We recall not the conclusion of Lovász Local Lemma (Lemma 1), but a crucial step in the proof (see [18], [13]): for any $\mathcal{T}$ with $i \notin \mathcal{T}$, we have $\Pr(A_i \mid \wedge_{j\in\mathcal{T}, j\neq i}\overline{A_j}) \leq x_i$, which in our case yields for any $M \in \mathcal{M}_i$

$$\Pr(A_M \mid \wedge_{M'\in\mathcal{T}, M'\neq M}\overline{A_{M'}}) \leq x_i = \frac{2p_{N,i}}{1 + \sqrt{1 - 8r\sum_{j\in I} d_j p_{N,j}}}.$$

Assume that $S = \{M_1, M_2, ..., M_s\}$. We have

$$\Pr(\overline{A_{M_1}} \wedge \overline{A_{M_2}} \wedge .... \wedge \overline{A_{M_s}} \mid \wedge_{M\in\mathcal{T}}\overline{A_M}) =$$

$$\prod_{\ell=1}^{s}\left[1 - \Pr\left(A_{M_\ell} \mid \overline{A_{M_1}} \wedge \overline{A_{M_2}} \wedge .... \wedge \overline{A_{M_{\ell-1}}} \wedge (\wedge_{M\in\mathcal{T}}\overline{A_M})\right)\right] \geq \prod_{\ell=1}^{s}(1 - x_{|M_\ell|}).$$

The proof of Lemma 5 is finished. Note that it also works with $\mathcal{T} = \emptyset$. $\qquad\square$

Lemma 6 *For any collection $\mathcal{M}$ of matchings in K_N, we have*

$$\Pr(\wedge_{M\in\mathcal{M}}\overline{A_M^{N+2}}) \leq \frac{2}{1 + \sqrt{1 - 8r\sum_{i\in I} d_i p_{N,i}}}\Pr(\wedge_{M\in\mathcal{M}}\overline{A_M^N}).$$

Proof: Partition Ω_{N+2}, introduce $\mathcal{C}_i$ and $\mathcal{B}_i$ as in the proof of Lemma 4, and use the fact derived there between (10) and (11) that

$$\Pr(\wedge_{M\in\mathcal{M}}\overline{A_M^{N+2}}) = \frac{1}{N+1}\sum_{i=1}^{N+1}\Pr(\wedge_{M\in\mathcal{B}_i}\overline{A_M^N}). \tag{16}$$

10

We are going to apply Lemma 5 with $S = \mathcal{M} \setminus \mathcal{B}_i$ and $\mathcal{T} = \mathcal{B}_i$. The number of matchings of size k in $\mathcal{M} \setminus \mathcal{B}_i$ is at most d_k. We have

$$
\begin{aligned}
\Pr(\wedge_{M \in \mathcal{M} \setminus \mathcal{B}_i} \overline{A_M^N} \mid \wedge_{M \in \mathcal{B}_i} \overline{A_M^N}) &\geq \prod_{k \in I} \left(1 - \frac{2 p_{N,k}}{1 + \sqrt{1 - 8r \sum_{j \in I} d_j p_{N,j}}} \right)^{d_k} \\
&\geq 1 - \sum_{k \in I} \frac{2 d_k p_{N,k}}{1 + \sqrt{1 - 8r \sum_{j \in I} d_j p_{N,j}}} \\
&= 1 - \frac{1 - \sqrt{1 - 8r \sum_{j \in I} d_j p_{N,j}}}{4r} \\
&\geq \frac{1 + \sqrt{1 - 8r \sum_{j \in I} d_j p_{N,j}}}{2}.
\end{aligned}
$$

We have from (16) and the estimate above:

$$
\begin{aligned}
\Pr(\wedge_{M \in \mathcal{M}} \overline{A_M^{N+2}}) &= \frac{1}{N+1} \sum_{i=1}^{N+1} \Pr(\wedge_{M \in \mathcal{B}_i} \overline{A_M^N}) \\
&\leq \frac{1}{N+1} \sum_{i=1}^{N+1} \Pr(\wedge_{M \in \mathcal{M}} \overline{A_M^N}) \frac{2}{1 + \sqrt{1 - 8r \sum_{j \in I} d_j p_{N,j}}} \\
&= \frac{2}{1 + \sqrt{1 - 8r \sum_{j \in I} d_j p_{N,j}}} \Pr(\wedge_{M \in \mathcal{M}} \overline{A_M^N}).
\end{aligned}
$$

The proof of Lemma 6 is finished. $\qquad\square$

Proof of Theorem 3: We are going to show that the negative dependency graph G defined for matchings of K_N in $\mathcal{M}$ is also an ϵ-near-positive dependency graph with ϵ as in (13); and then Theorem 1 together with (13) will finish the proof of (14) and Theorem 3 The first part of the definition, $\Pr(A_i \wedge A_j) = 0$ for ij edges comes for free. We focus on the second part.

For any $F \in \mathcal{M}_i$ and a subset $S \subseteq \overline{N_G(F)}$, we need to prove

$$
\Pr(A_F \mid \wedge_{M \in S} \overline{A_M}) \geq (1 - \epsilon) \Pr(A_F),
$$

or equivalently,

$$
\Pr(\wedge_{M \in S} \overline{A_M} \mid A_F) \geq (1 - \epsilon) \Pr(\wedge_{M \in S} \overline{A_M}).
$$

Let $S' = \{M \setminus F \mid M \in S\}$. Observe that $\emptyset \notin S'$. Note that

$$
\begin{aligned}
\Pr(\wedge_{M \in S} \overline{A_M} \mid A_F) &= \frac{\Pr(\wedge_{M \in S} \overline{A_M} \wedge A_F)}{\Pr(A_F)} \quad (17) \\
&= \frac{\Pr(\wedge_{M \in S} \overline{A_{M \setminus F}} \wedge A_F)}{\Pr(A_F)} \\
&= \Pr(\wedge_{M \in S'} \overline{A_M} \mid A_F). \quad (18)
\end{aligned}
$$

11

We have

$$\Pr(\wedge_{M\in S'}\overline{A_M} \mid A_F) = \Pr(\wedge_{M\in S'}\overline{A_M^{N-2i}}) \tag{19}$$

$$= \Pr(\wedge_{M\in S'}\overline{A_M^N}) \prod_{j=1}^{i} \frac{\Pr(\wedge_{M\in S'}\overline{A_M^{N-2j}})}{\Pr(\wedge_{M\in S'}\overline{A_M^{N-2j+2}})}$$

$$(\text{by Lemma 6}) \geq \Pr(\wedge_{M\in S'}\overline{A_M^N}) \prod_{\ell=0}^{i-1} \frac{1+\sqrt{1-8r\sum_{j\in I}d_j p_{N-2\ell,j}}}{2}$$

$$\geq \Pr(\wedge_{M\in S'}\overline{A_M^N}) \prod_{\ell=0}^{r-1} \frac{1+\sqrt{1-8r\sum_{j\in I}d_j p_{N-2\ell,j}}}{2} \tag{20}$$

For any M, which does not conflict to F, we have $\overline{A_{M\setminus F}} \subset \overline{A_M}$. We have with $S' = \{M\setminus F \mid M \in S\}$ that

$$\frac{\Pr(\wedge_{M\in S'}\overline{A_M^N})}{\Pr(\wedge_{M\in S}\overline{A_M^N})} = \frac{\Pr(\wedge_{M\in S}\overline{A_{M\setminus F}^N})}{\Pr(\wedge_{M\in S}\overline{A_M^N})} \tag{21}$$

$$= \frac{\Pr(\wedge_{M\in S}\overline{A_{M\setminus F}^N} \wedge \overline{A_M^N})}{\Pr(\wedge_{M\in S}\overline{A_M^N})}$$

$$= \frac{\Pr([\wedge_{M\in S, M\cap F\neq\emptyset}\overline{A_{M\setminus F}^N}] \wedge [\wedge_{M\in S}\overline{A_M^N}])}{\Pr(\wedge_{M\in S}\overline{A_M^N})}$$

$$= \Pr(\wedge_{M\in S'\setminus S}\overline{A_M^N} \mid \wedge_{M\in S}\overline{A_M^N}). \tag{22}$$

Now apply Lemma 5 to $S'\setminus S$, S and $S\cup S'$ instead of S, $\mathcal{T}$ and $\mathcal{M}$. To have positive number under the squareroot in the formula corresponding to formula (15), we need only that

$$8r\sum_{j\in I(\mathcal{M}\cup\mathcal{M}_F)}d_j(\mathcal{M}\cup\mathcal{M}_F)p_{N,j}$$

$$\leq 8r\sum_{j\in I(\mathcal{M})}d_j(\mathcal{M})p_{N,j} + 8r\sum_{j\in I(\mathcal{M}_F)}d_j(\mathcal{M}_F)p_{N,j}$$

$$\leq (1-8\delta) + \sum_{M'\in\mathcal{M}, M'\cap F\neq\emptyset}p_{N,|M'\setminus F|} \leq 1-7\delta$$

by the 2. and 3. conditions of δ-sparseness. We have

$$\Pr(\wedge_{M\in S'\setminus S}\overline{A_M^N} \mid \wedge_{M\in S}\overline{A_M^N})$$

$$\geq \prod_{M\in S'\setminus S}\left(1 - \frac{2p_{N,|M|}}{1 + \sqrt{1 - 8r\sum_{j\in I(\mathcal{M}\cup\mathcal{M}_F)}d_j(\mathcal{M}\cup\mathcal{M}_F)p_{N,j}}}\right)$$

$$\geq \prod_{M\in S'\setminus S}(1 - 2p_{N,|M|})$$

$$\geq 1 - \sum_{M\in S'\setminus S}2p_{N,|M|}$$

$$\geq 1 - 2\delta. \tag{23}$$

Finally, we have

$$
\begin{aligned}
&\Pr(\wedge_{M\in S}\overline{A_M} \mid A_F) \\
\text{by (17-18)} \quad =\quad & \Pr(\wedge_{M\in S'}\overline{A_M} \mid A_F) \\
\text{by (19-20)} \quad \geq\quad & \Pr(\wedge_{M\in S'}\overline{A_M^N})\prod_{i=0}^{r-1}\frac{1 + \sqrt{1 - 8r\sum_{j\in I}d_jp_{N-2i,j}}}{2} \\
\text{by (21-22)} \quad =\quad & \Pr(\wedge_{M\in S}\overline{A_M^N})\Pr(\wedge_{M\in S'\setminus S}\overline{A_M^N} \mid \wedge_{M\in S}\overline{A_M^N}) \\
& \times \prod_{i=0}^{r-1}\frac{1 + \sqrt{1 - 8r\sum_{j\in I}d_jp_{N-2i,j}}}{2} \\
\text{by (23)} \quad \geq\quad & \Pr(\wedge_{M\in S}\overline{A_M^N})(1 - 2\delta)\prod_{i=0}^{r-1}\frac{1 + \sqrt{1 - 8r\sum_{j\in I}d_jp_{N-2i,j}}}{2}.
\end{aligned}
$$

Thus, the negative dependency graph G is also a ϵ-positive dependency graph. The proof is finished by Theorem 1. $\qquad\square$

5 Asymptotic results in the matching model

5.1 The key result

A collection $\mathcal{M}$ of matchings of K_N or $K_{N,N}$ is called *regular* if for every i there is an integer d_i so that for any vertex v the number of machings in $\mathcal{M}$ containing v is d_i, independent of v.

Theorem 4 *Suppose $\mathcal{M}$ is a regular collection of matchings of K_N or $K_{N,N}$. Let $\mu = \sum_{M\in\mathcal{M}}\Pr(A_M)$. Suppose $\mu = o(\sqrt{N}r^{-3/2})$ but μ is separated from zero; and also that $\mathcal{M}$ is δ-sparse for some $\delta = o(\mu^{-1})$. Then we have*

$$\Pr(\wedge_{M\in\mathcal{M}}\overline{A_M}) = (1 + o(1))e^{-\mu}.$$

Proof: We produce the proof for K_N and leave the almost identical proof for $K_{N,N}$ to the reader. Since $\mathcal{M}$ is regular, for each $i \in I$, we have

$$d_i = \frac{2i}{N}|\mathcal{M}_i|.$$

As $\mu = \sum_{i \in I}|\mathcal{M}_i|p_{N,i}$, we have

$$\sum_{i \in I} d_i p_{N,i} = \sum_{i \in I}|\mathcal{M}_i|p_{N,i}\frac{2i}{N} \leq \frac{2r\mu}{N}. \tag{24}$$

We will use the inequality $\frac{1}{1+\sqrt{1-x}} \leq 1+x$ and the facts that $\mu = o(\sqrt{N}r^{-3/2})$ imply $\mu = o(N/r^2)$, $\mu = o(N)$, $r^2\mu^2/N = o(1)$, which in turn implies $r^2/N = o(1)$ as μ is separated from zero. To set the lower bound, Lemma 5 with $S = \mathcal{M}$ and $\mathcal{T} = \emptyset$ implies

$$
\begin{aligned}
\Pr(\wedge_{M \in \mathcal{M}}\overline{A_M}) \;\geq\; & \prod_{i \in I}\left(1 - \frac{2p_{N,i}}{1 + \sqrt{1 - 8r\sum_{j \in I}d_j p_{N,j}}}\right)^{|\mathcal{M}_i|} \\
\geq\; & \prod_{i \in I}\left(1 - \frac{2p_{N,i}}{1 + \sqrt{1 - 8r\frac{2r\mu}{N}}}\right)^{|\mathcal{M}_i|} \\
=\; & \prod_{i \in I}\left(1 - [1 + 16r^2\frac{\mu}{N})]p_{N,i}\right)^{|\mathcal{M}_i|} \\
& (\text{using } 1 - x = e^{\ln(1-x)} = e^{-x - O(x^2)} \text{ and } p_{N,i} \leq 1/N) \\
=\; & e^{-\mu - 16r^2\frac{\mu^2}{N} + O(\frac{\mu}{N})} \\
=\; & (1 - o(1))e^{-\mu}.
\end{aligned}
$$

To set the upper bound, recall Theorem 3 to find a δ-near positive dependency graph. Using (24) and the fact that $P_{N-2k,i} = [1 + O(r^2/N)]p_{N,i}$ follows from $r^2/N = o(1)$, we obtain

$$
\begin{aligned}
\prod_{k=1}^{r}\frac{1 + \sqrt{1 - 8r\sum_{i \in I}d_i p_{N-2k,i}}}{2} \;=\; & \prod_{k=1}^{r}\frac{1 + \sqrt{1 - 8r\sum_{i \in I}(1 + O(\frac{r^2}{N}))d_i p_{N,i}}}{2} \\
\geq\; & \prod_{k=1}^{r}\frac{1 + \sqrt{1 - \frac{16r^2\mu}{N}[1 + O(\frac{r^2}{N})]}}{2} \\
\geq\; & \prod_{k=1}^{r}\left(1 - \frac{4r^2\mu}{N} - O(\frac{r^4\mu^2}{N^2})\right) \\
\geq\; & 1 - \frac{4r^3\mu}{N} - O(\frac{r^5\mu^2}{N^2}).
\end{aligned}
$$

14

(Note that in the argument above, we actually proved the 2. condition of δ-sparseness, from the other 3 conditions and the conditions for μ in Theorem 4.) Thus,

$$
\begin{aligned}
\Pr(\wedge_{M \in \mathcal{M}} \overline{A_M}) & \leq \prod_{M \in \mathcal{M}} \left(1 - (1 - 2\delta) \prod_{k=1}^{r} \frac{1 + \sqrt{1 - 8r \sum_{i \in I} d_i p_{N-2k,i}}}{2} \Pr(A_M) \right) \\
& \leq \prod_{M \in \mathcal{M}} \left(1 - (1 - 2\delta) \left[1 - \frac{8r^3 \mu}{N} - O(\frac{r^5 \mu^2}{N^2}) \right] \Pr(A_M) \right) \\
& \leq e^{-(1-2\delta)(\mu - \frac{8r^3 \mu^2}{N} + O(\frac{r^5 \mu^3}{N^2}))} \\
& = (1 + o(1)) e^{-\mu} \quad \text{(here we used } \delta = o(\mu^{-1})).
\end{aligned}
$$

We need uniform $O()$ estimates for these calculations, but those are easy to obtain. $\qquad \square$

5.2 Applications I: Counting k-cycle free permutations and Latin rectangles

It is known and easy that for any fixed k, the probability of a random permutation not having any k-cycle is asymptotically $e^{-1/k}$, see Borkowitz [6]. In our earlier paper, [14], we obtained an $(1 - o(1)) e^{-1/k}$ lower bound for the probability from Lovász Local Lemma. Now we show that machinery that we developed in this paper actually yields the asymptotic formula.

Let us be given two N-element sets with elements $\{1, 2, ..., N\}$ and $\{1', 2', ..., N'\}$. Let us identify a permutation of the first set, π, with a mathching between the two sets, such that i is joined to $\pi(i)'$. A k-cycle in the permutation can be identified as a matching between $K \subset \{1, 2, ..., N\}$ to $\{\ell' : \ell \in K\}$ with $|K| = k$, which does not have a proper nonempty subset with the same property. The bad events for the negative dependency graph are these k-element matchings, there are $\binom{N}{k}(k-1)!$ of them. Hence $\mathcal{M} = \mathcal{M}_k$ and $r = k$. These collection of matchings is regular with $d = \binom{N}{k}(k-1)!/N = (N)_k/(kN)$. We have $\mu = \binom{N}{k}(k-1)!(N-k)!/N! = \frac{1}{k}$ and $p_{N,k} = \frac{1}{(N)_k}$. For the conditions of δ-sparseness, 1. is trivial, 2. follows from $\frac{(N)_k}{kN} \frac{1}{(N-2k)_k} = o(1/N)$, as for 3. we will select a $\delta = O(1/N)$, which will also take care of 4.

We are left with justifying condition 3. Fix a k-matching F corresponding to a k-cycle in the permutation. Take j edges of F with $1 \leq j \leq k-1$. These edges can be selected $\binom{k}{j}$ ways which is bounded. The j edges cover some s numbers from $\{1, 2, ..., N\}$, either with prime or not. As the j edges make just a part of the cycle, $j + 1 \leq s$. An edge set M' in the complete bipartite graph, which corresponds to different k-cycle, uses $k - s$ additional numbers from $\{1, 2, ..., N\}$, both with and without prime. Those additional vertices can be selected in $\binom{N-s}{k-s}$ ways. There is a bounded number of ways to complete the cycle of M' as the vertex set is given. Now $p_{N, |M' \setminus F|} = \frac{1}{(N)_{k-j}}$,

and $\sum_{M \in \mathcal{M}_F} p_{N,|M|} \le c_k \binom{N-(j+1)}{k-(j+1)} = O(1/N)$.

The additional conditions of Theorem 4 also hold, as μ is separated from zero, in fact constant, and $\delta = o(1)$. We also have $\mu = \frac{1}{k} = o(\sqrt{N}k^{-3/2})$. Therefore Theorem 4 applies, and the number of k-cycle free permutations is $(1 + o(1))e^{-1/k}$.

Let us turn now to the enumeration of Latin rectangles. A $k \times n$ Latin rectangle is a sequence of k permutations of $\{1, 2, ..., n\}$ written in a matrix form, such that no column has any repeated entries. Let $L(k, n)$ denote the number of $k \times n$ Latin rectangles. $L(2, n)$ is just $n!$ times the number of derangements, i.e. $(n!)^2 e^{-1}$. In 1944, Riordan [17] showed that $L(3, n) \sim (n!)^3 e^{-2}$. In 1946, Erdős and Kaplansky [9] showed

$$L(k, n) \sim (n!)^k e^{-\binom{k}{2}}$$

for $k = o((\log n)^{3/2})$. In 1951, Yamamoto [21] extended this asymptotic formula for $k = o(n^{1/3})$. In 1978, Stein [20] refined the asymptotic formula to

$$L(k, n) \sim (n!)^k e^{-\binom{k}{2} - \frac{k^3}{6n}} \tag{25}$$

using the Chen-Stein method [7], and extended the range to $k = o(n^{1/2})$. The current best asymptotic formula is due to Godsil and McKay [12], whose further refined formula works for $k = o(n^{6/7})$.

Formula (25) has had an unexpected proof by Skau [19], who proved the inequality

$$(n!)^k \prod_{t=1}^{k-1} \left(1 - \frac{t}{n}\right)^n \le L(k, n), \tag{26}$$

from the van der Waerden inequality for the permanent, providing the lower bound for the asymptotic formula. The upper bound followed from Minc's inequality for the permanent. In [14] we derived (26) from Lemma 1 for $k = o((n/\log n)^{1/2})$.

Fix an arbitrary $t \times n$ Latin rectangle now with rows $\pi_1, \pi_2, ..., \pi_t$. Define the event A_{ij} by $\pi_i(j) = \pi_{t+1}(j)$, in other words, we look for matchings from $\{1, 2, ..., n\}$ to $\{1', 2', ..., n'\}$ such that the one-edge matchings $(\pi_{t+1}(j), \pi_i(j)')$ are to be avoided. In other words, $\mathcal{M}^{(t)} = \{(\pi_{t+1}(j), \pi_i(j)') : \pi_i(j) = \pi_{t+1}(j)\}$. These are canonical events. Let G_t be the graph whose vertices are the (i, j) entries for $j = 1, 2, ..., n$, $i = 1, 2, ..., t$, and every (i_1, j) is joined with every (i_2, j). The maximum degree in this graph is $t - 1 = o(n^{1/2})$. With the choice $x_{ij} = 2/n$ these events satisfy (3) in the graph G, and therefore the graph G according to Theorem 2 is a negative dependency graph. Define the events $B_j = \vee_{1 \le i \le t} A_{ij}$. Clearly $\Pr(B_j) \le t/n$. Applying the first part of Lemma 3, the quotient graph is empty, and is a negative dependency graph for the B_j events. Lemma 1 applies and $\Pr(\wedge_{j=1}^n \overline{B_j}) \ge (1 - t/n)^n$. Hence we have

$$n! \left(1 - \frac{t}{n}\right)^n \le \frac{L(t+1, n)}{L(t, n)} \tag{27}$$

as we did in [14]. Iterating this estimate, formula (26) follows.

Our next goal is to prove an upper bound corresponding to the right hand side of (25), in addition to the lower bound, using our method. Accept for the time being that $\mathcal{M}_t$ and the single-edge is δ_t-sparse. Then we obtain from Theorem 3 (13) that $\mathcal{M}_t$ is also ϵ_t-near-positive dependency graph, where (13) defines ϵ_t with $r(\mathcal{M}_t) = 1$, $\delta_t(\mathcal{M}_t)$, $d_1(\mathcal{M}_t)$. It follows from the second part of Lemma 3 that

$$\frac{L(t+1,n)}{L(t,n)} \leq n!\left(1 - (1-\epsilon_t)\frac{t}{n}\right)^n,\tag{28}$$

and iterating (28), we obtain

$$L(k,n) \leq (n!)^k \prod_{t=1}^{k-1}\left(1 - (1-\epsilon_t)\frac{t}{n}\right)^n.\tag{29}$$

We claim that $\mathcal{M}_t$ is $\delta_t = 0$-sparse, as the summation in condition 3. of Theorem 3 is empty. Condition 1 and 4 are trivial. Condition 2. simplifies to $d_1(\mathcal{M}_t)p_{n,1} < 1/8$. As $p_{n,1} = 1/n$ and $d_1(\mathcal{M}_t) \leq t$, we have condition 2. Furthermore, from (13) we have $\epsilon_t = \frac{1}{2}(1-\sqrt{1-8d_1(\mathcal{M}_t)/n}) \leq \frac{1}{2}(1-\sqrt{1-8t/n}) \leq 4t/n$, and hence

$$L(k,n) \leq (n!)^k \prod_{t=1}^{k-1}\left(1 - \frac{t}{n} + 4\frac{t^2}{n^2}\right)^n.\tag{30}$$

Comparing the fraction of the upper bound in (30) and the lower bound in (26), we conclude that the upper bound is asymptotically tight for $k = o(n^{1/3})$:

$$\prod_{t=1}^{k-1}\left[1 + 4\frac{t^2}{n^2}(1 - \frac{t}{n})\right] \leq \exp\left(\sum_{t=1}^{k-1} 4\frac{t^2}{n^2}(1 - \frac{t}{n})\right).$$

5.3 Applications II: The configuration model and the enumeration of d-regular graphs

For a given sequence of positive integers with an even sum, $\mathbf{d} = (d_1, d_2, \ldots, d_n)$, the *configuration model of random multigraphs with degree sequence* $\mathbf{d}$ is defined as follows [4].

1. Let us be given a set U that contains $N = \sum_{i=1}^n d_i$ distinct mini-vertices. Let U be partitioned into n classes such that the ith class consists of d_i mini-vertices. This ith class will be associated with vertex v_i after identifying its elements through a *projection*.

2. Choose a random matching M of the mini-vertices in U uniformly.

3. Define a random multigraph G associated with M as follows: For any two (not necessarily distinct) vertices v_i and v_j, the number of edges joining v_i and v_j in G is equal to the total number of edges in M between mini-vertices associated with v_i and mini-vertices associated with v_j.

The configuration model of random d-regular graphs on n vertices is the instance $d_1 = d_2 = \cdots = d_n$, where nd is even.

The enumeration problem of labelled d-regular graphs has a rich history in the literature. The first result was Bender and Canfield [3], who showed in 1978 that for any fixed d, with nd even, the number of them is

$$\sqrt{2}e^{(1-d^2)/4}\left(\frac{d^d n^d}{e^d(d!)^2}\right)^{\frac{n}{2}}.$$

The same result was discovered at the same time by Wormald. In 1980, Bollobás [4] introduced probability to this enumeration problem by defining the configuration model, and put the result in the alternative form

$$(1+o(1))e^{\frac{1-d^2}{4}}\frac{(dn)!}{(dn/2)!2^{dn/2}(d!)^n}. \tag{31}$$

where the term $(1+o(1))e^{\frac{1-d^2}{4}}$ in (31) can be explained as the probability of obtaining a simple graph after the projection. The term $\frac{(dn)!}{(dn/2)!2^{dn/2}}$ equals $(dn-1)!!$, the number of perfect matchings on dn elements, and $\frac{1}{(d!)^n}$ is just the number of ways matchings can yield the same simple graph after projection. Bollobás also extended the range of the asymptotoc formula to $d < \sqrt{2\log n}$, which was further extended to $d = o(n^{1/3})$ by McKay [15] in 1985. The strongest result is due to McKay and Wormald [16] in 1991, who refined the probability of obtaining a simple graph after the projection to $(1+o(1))e^{\frac{1-d^2}{4}-\frac{d^3}{12n}+O(\frac{d^2}{n})}$ and extended the range of the asymptotic formula to $d = o(n^{1/2})$. Wormald's Theorem 2.12 in [23] (originally published in [22] asserts that for any fixed numbers $d \geq 3$ and $g \geq 3$, the number of labelled d-regular graphs with girth at least g, is

$$(1+o(1))e^{-\sum_{i=1}^{g-1}\frac{(d-1)^i}{2i}}\frac{(dn)!}{(dn/2)!2^{dn/2}(d!)^n}. \tag{32}$$

In our theorem below, we allow both d and g go to infinity slowly. If we set $g = 3$, we get back asymptotic formula for the number of d-regular graphs up to $d = o(n^{1/3})$, giving an alternative proof to McKay's result cited above. However, our method inherently fail to extend this result as McKay and Wormald did see above). In fact, our method fails to extend the lower bound in this way.

Theorem 5 *In the configuration model, if $d \geq 3$ and*

$$g^3 d^{2g-3} = o(n), \tag{33}$$

then the probability that the random d-regular multigraph has girth at least $g \geq 3$ is

$$(1+o(1))e^{-\sum_{i=1}^{g-1}\frac{(d-1)^i}{2i}}.$$

The case $g = 3$ means that the random d-regular multigraph is actually a simple graph.

Proof: For $i = 1, 2, \ldots, g - 1$, let $\mathcal{M}_i$ be the set of matchings of U whose projection gives a cycle of length i. Observe that $\mathcal{M} = \cup_{i=1}^{g-1} \mathcal{M}_i$ is regular. We have $d_1 = d - 1$, $d_2 = (n-1)d(d-1)^2$, and for $i \geq 3$,

$$d_i = (n-1)(n-2)\ldots(n-i+1)d^{i-1}(d-1)^i.$$

$$
\begin{aligned}
\mu &= \sum_{i \in I} |\mathcal{M}_i| p_{N,i} = \sum_{i=1}^{g-1} \frac{N}{2i} d_i p_{N,i} \\
&= \sum_{i=1}^{g-1} \frac{N}{2i} \cdot \frac{(n-1)(n-2)\ldots(n-i+1)d^{i-1}(d-1)^i}{(N-1)(N-3)\cdots(N-2i+1)} \\
&= (1 + o(1)) \sum_{i=1}^{g-1} \frac{(d-1)^i}{2i},
\end{aligned}
$$

where the $o(1)$ estimate needs $g = o(\sqrt{n})$, which follows from the condition (33).

We have to verify that $\mathcal{M}$ is δ-sparse for some $\delta = o(\mu^{-1})$. Condition 1. holds: no matching from $\mathcal{M}$ is a subset of another matching from $\mathcal{M}$. For if it happened, after the projection we would have the edge set of a cycle as a subset of the edge set of another cycle.

Condition 2. holds: for $r = g - 1$, we find that the expected

$$8(g-1) \sum_{i=1}^{g-1} d_i p_{N-2g,i} = o(1)$$

is equivalent to $gd^{g-2} = o(n)$, which follows from (33). We will set a δ later satisfying 4., which will also be $o(1)$.

Condition 3. holds: For any $F \in \mathcal{M}$, we need estimate $\sum_{M \in \mathcal{M}_F} p_{N,|M|}$. If the projection of F is a loop, then $\mathcal{M}_F = \emptyset$. Now we assume the projection of F is a cycle C_k. Assume that $M' \in \mathcal{M}$ intersects F, $M = M' \setminus F$, and the projection of M' is a cycle C_l with $k, l \leq g - 1$. Then the components of $C_l \cap C_k$ are $t \geq 1$ paths $P_1, P_2, \ldots, P_t$. Fixing the paths, and the edges in $M' \cap F$, some additional ℓ vertices are joined with these t paths to make C_l. So the number of C_l's (i.e. M' matchings) with fixed paths and $M' \cap F$, is at most

$$\sum_{\ell \leq g-1-2t} \binom{n}{\ell} (\ell + t - 1)! 2^t,$$

and the M'-s is at most $d^{\ell+2t}$ times more. The paths can be selected $\sum_{t=1}^{k/2} \binom{k}{2t}$ ways. The associated $p_{N,|M' \setminus F|}$ is at most $(N - 2g)^{-(\ell+t)}$. We summarize that

$$\sum_{M \in \mathcal{M}_F} p_{N,|M|} \leq \sum_{t=1}^{k/2} \binom{k}{2t} \sum_{\ell \leq g-1-2t} \binom{n}{\ell} (\ell + t - 1)! 2^t \frac{d^t}{n^{\ell+t}}. \tag{34}$$

19

As $\ell + t - 1 \le g - 3$, we have $(\ell + t - 1)_{t-1} \le (g-2)_{t-1}$, and (34) is at most

$$\sum_{t=1}^{k/2} \binom{k}{2t}\left(\frac{2d(g-2)}{n}\right)^t = O(dg^4/n), \text{ as the last sum has its biggest term at } t = 1.$$

Take $\delta = Kdg^4/n$ with a large constant and condition 3. has been verified. Condition 4 and $\delta = o(\mu^{-1})$ both follow from (33). Finally, the condition $\mu = o(\sqrt{N}r^{-3/2})$ of Theorem 4, with $r = g - 1$ and $N = nd$, follows from We have $\mu = \Theta(d^{g-1}) = o\big((nd)^{1/2}g^{-3/2}\big)$ from the condition (33) again. μ is separeted from zero. Therefore Theorem 4 applies. $\qquad\square$

6 Revisiting girth and chromatic number

6.1 Some technical lemmas about matchings

This section proves two technical lemmas about matchings that we will use in the last Section. Let N be an even positive integer. For a set $S \subset [N]$, we say that a perfect matching M of K_N *traverses* S, if every edge in M is incident to at most one vertex in S, in other words no edge has two endpoints in S.

Lemma 7 *For a fixed set S of size s, the probability that S is traversed, equals to*

$$\frac{2^s \binom{\frac{N}{2}}{s}}{\binom{N}{s}}.$$

Proof: Clearly the probability in question does not depend on the choice of S, just depends on the cardinality s. Therefore the probability does not change if we average it out for all s-subsets, and hence it is

$$\frac{\#(S, M) : \text{perfect matching } M \text{ traverses } S}{(N-1)!! \binom{N}{s}}.$$

Count now in the ordered pairs in the numerator as follows: for all $(N-1)!!$ perfect matchings, decide which s edges of the $N/2$ edges of the perfect matching have endpoint in S, and for those s edges decide which endpoint out of the two possibilities will belong to S. $\qquad\square$

Lemma 8 *Assume that $\frac{\ln^2 N}{N^{1/3}} \le x \le \frac{1}{4}$ and $xN \to \infty$. For any fixed set S of size xN, the probability that S is traversed is*

$$e^{-N\frac{x^2}{2}+O(Nx^3)},$$

where $O()$ refers to $xN \to \infty$.

Proof: ¿From the Stirling formula

$$N! = \left(\sqrt{2\pi N} + O(N^{-\frac{1}{2}})\right)\frac{N^N}{e^N}$$

20

one easily obtains

$$\binom{N}{xN} = \frac{1 + O(\frac{1}{xN})}{\sqrt{2\pi x(1-x)N}} e^{N \cdot H(x)}.$$

Here $H(x) = -x \ln x - (1-x)\ln(1-x)$ denotes the binary entropy function. Also,

$$\binom{\frac{N}{2}}{xN} = \frac{1 + O(\frac{1}{xN})}{\sqrt{2\pi x(1-2x)N}} e^{\frac{1}{2}N \cdot H(2x)},$$

and finally we have

$$
\begin{aligned}
\frac{2^{xN}\binom{\frac{N}{2}}{xN}}{\binom{N}{xN}} &= \left(1 + O(\frac{1}{xN})\right) \frac{2^{xN} \frac{1}{\sqrt{2\pi x(1-2x)N}} e^{\frac{1}{2}N \cdot H(2x)}}{\frac{1}{\sqrt{2\pi x(1-x)N}} e^{N \cdot H(x)}} \\
&= \left(1 + O(\frac{1}{xN})\right) \sqrt{\frac{1-x}{1-2x}} e^{N(\frac{1}{2}H(2x) - H(x) + x \ln 2)} \\
&= \left(1 + O(\frac{1}{xN})\right) \sqrt{\frac{1-x}{1-2x}} e^{-N\left((1/2-x)\ln(1-2x) - (1-x)\ln(1-x)\right)} \\
&= e^{-Nx^2/2 + O(x^3 N)},
\end{aligned}
$$

where the last inequality follows from $x = \Omega(\frac{\ln^2 N}{N^{1/3}})$. $\qquad\square$

6.2 High girth and high chromatic number graphs on a given degree sequence

We continue with further definitions for the configuration model, using the notation from Section 5.3. For any subset S of $V(G)$, we define the volume of S

$$Vol(S) = \sum_{i \in S} d_i.$$

Let $\Delta = \max_i d_i$ denote the maximum degree in G; and let $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$ denote the *average degree*, and $\tilde{d} = \sum_{i=1}^{n} d_i^2 / \sum_{i=1}^{n} d_i$ the *second order average degree* in G. We have

$$\bar{d} \le \tilde{d} \le \Delta,$$

and any of the inequalities holds with equality if and only if all degrees are equal.

Recall that the $girth(G)$ of a graph G is the size of its shortest cycle, and $\chi(G)$ is the chromatic number of G.

An early result of Erdős [8] asserts that for every k and l, there is a graph G with $girth(G) > l$ and $\chi(G) \ge k$. In the following theorem we refine this result of Erdős by providing such a graph G, which in addition, has a prescribed degree sequence satisfying some mild conditions.

Theorem 6 *There exists an ℓ_0, such that for every k and $\ell \geq \ell_0$, and every sequence of positive integers with even sum, $\mathbf{d} = (d_1, d_2, \ldots, d_n)$ that satisfies*

$$\tilde{d} \geq 3 \tag{35}$$

$$8k^2(\tilde{d}-1)^\ell < N \tag{36}$$

$$4\ell\Delta(\tilde{d}-1)^{\ell-1} < N/10, \tag{37}$$

there is a simple graph G, whose degree sequence $\mathbf{d}$, such that $girth(G) > \ell$ and $\chi(G) \geq k$. In particular, sequences satisfying (35), (36), and (37) are degree sequences. Furthermore, regular graphs with $d \geq 3$, satisfying (36), and (37), have $girth(G) > \ell$ and $\chi(G) \geq k$.

Proof: First of all, (35) and (37) imply that $\ell < \log_2 N$, and from here $\lim_{\ell \to \infty} N^\ell/(N-2\ell+1)^\ell = 1$.

We apply Lemma 2 to find a lower bound for the probability that G contains no short cycles. Bad events are of the form A_M where M is a fixed matching in U of r edges, such that such that under the projection from U to $V(G)$, the edges of M turn into a cycle $i_1 i_2 \cdots i_r$ of length r for some $1 \leq r \leq l$. (Cycles of length $r = 1$ or 2 are corresponding to loops or multiple edges in G.) It is obvious that $\Pr(A_m) = \frac{1}{(N-1)(N-3)(N-5)\cdots(N-2r+1)}$; and we have the estimates $\frac{1}{(N-2r+1)^r} \geq \Pr(A_m) \geq \frac{1}{(N-r)^r}$, where the last estimate follows from the inequality of arithmetic and geometric means.

Set $\epsilon = \frac{4\ell\Delta}{N} \frac{(\tilde{d}-1)^{\ell-1}N^\ell}{(N-2\ell+1)^\ell}$. For $\ell > \ell_0$, by (37), we have $\epsilon < 1/8$ as required in Lemma 2. To apply Lemma 2, we have to verify $\Pr(A_M) \leq \epsilon$, and

$$\sum_{M' \text{ conflicts to } M} \Pr(A_{M'}) \leq \epsilon.$$

As every $\Pr(A_M)$ is equal, the latter inequality implies the first.

Let us denote by $E(i_1 i_2 \ldots i_r)$ the union of the bad events M, each with r edges, all of which yield in G a particular $i_1 i_2 \ldots i_r$ r-cycle. To estimate $\Pr(E(i_1 i_2 \ldots i_r))$, select two elements in order from the i_jth class for $j = 1, 2, \ldots, r$, and join the second element selected from the i_jth class to the first element selected from the i_{j+1}th class, for every j, identifying index $r+1$ with 1. Now we have

$$
\begin{aligned}
\Pr(E(i_1 i_2 \ldots i_r)) &\leq \frac{\prod_{j=1}^r d_{i_j}(d_{i_j}-1)}{(N-1)(N-3)\ldots(N-2r+1)} \\
&\leq \frac{\prod_{j=1}^r (d_{i_j}^2 - d_{i_j})}{(N-2r+1)^r}.
\end{aligned}
$$

(The first inequality actually holds with equality for $r \geq 3$.) In order to use

22

Lemma 2, let us sum up first the probabilities of *all* bad events:

$$\sum_M \Pr(A_M) \ \leq \ \sum_{r=1}^{\ell} \sum_{\{i_1,i_2,\ldots,i_r\}} (r-1)! \Pr(E(i_1 i_2 \ldots i_r))$$

$$= \ \sum_{r=1}^{\ell} \sum_{\{i_1,i_2,\ldots,i_r\}} (r-1)! \frac{\prod_{j=1}^{r}(d_{i_j}^2 - d_{i_j})}{(N-2r+1)^r}$$

$$= \ \sum_{r=1}^{\ell} \frac{1}{r} \frac{\left(\sum_i (d_i^2 - d_i)\right)^r}{(N-2r+1)^r}$$

$$= \ \sum_{r=1}^{\ell} \frac{1}{r} \frac{(\tilde{d}-1)^r N^r}{(N-2r+1)^r}$$

$$< \ \frac{2(\tilde{d}-1)^{\ell} N^{\ell}}{(N-2\ell+1)^{\ell}}.$$

Let M be now the bad event defined by some r edges in a matching of U, and let M' be a bad event defined by some s edges in a matching of U, such that M' is in conflict with M. How many M' can be there? Assume that M is yielding the cycle $i_1 i_2 ... i_r$ in G. The s edges of M' must share at least one vertex with the r edges of M, say one of the vertices from the ith class. There are 2 choices for this vertex, and then $d_t - 1$ choices for the other vertex of M' in the tth class. One selects further $j_2, j_3, ..., j_s$ classes, puts them and t in order, and proceeds like we did at the estimation for the event $E(i_1 i_2 ... i_r)$. Therefore, the number of M' conflicting events is at most

$$\left(2\sum_{t=1}^{r} d_t\right)(s-1)! \sum_{\{j_2,j_3,\ldots,j_s\}} \prod_{u=2}^{s} (d_{j_u}^2 - d_{j_u}).$$

Like in the calculations before,

$$\sum_{M' \text{ conflicts to } M} \Pr(A_{M'}) \ \leq \ \sum_{s=1}^{\ell} 2r\Delta(s-1)! \sum_{\{j_2,j_3,\ldots,j_s\}} \frac{\prod_{u=2}^{s}(d_{j_u}^2 - d_{j_u})}{(N-2s+1)^s}$$

$$= \ 2r\Delta \sum_{s=1}^{\ell} \frac{\left(\sum_i (d_i^2 - d_i)\right)^{s-1}}{(N-2s+1)^s}$$

$$< \ \frac{4\ell\Delta}{N} \frac{(\tilde{d}-1)^{\ell-1} N^{\ell}}{(N-2\ell+1)^{\ell}} = \epsilon.$$

By Lemma 2, the probability that $girth(G) > \ell$ is at least

$$\exp\left(-(1+2\epsilon)\frac{2(\tilde{d}-1)^{\ell} N^{\ell}}{(N-2\ell+1)^{\ell}}\right). \tag{38}$$

Now we set an upper bound on the probability that G is k-colorable. If G is k-colorable, then G contains an independent set of volume at least $\frac{N}{k}$. By Lemma 8 at $x = 1/k$, the probability of this event is at most

$$2^n \exp\left(-\frac{N}{2k^2} + O(\frac{N}{k^3})\right) = \exp\left(-\frac{N}{2k^2} - (\ln 2)/\bar{d} + O(\frac{1}{k^3})\right). \qquad (39)$$

Now it follows from condition (36) that the probability of k-colorability in (39) is less than probability of $girth(G) > \ell$ in (38), and therefore with positive probability, G has girth greater than ℓ and has chromatic number greater than k. $\qquad \square$

References

[1] N. Alon, J. H. Spencer, *The Probabilistic Method,* second edition, John Wiley and Sons, New York, 2000.

[2] J. Beck, An algorithmic approach to the Lovász local lemma. I., *Random Structures and Algorithms* **2** (1991) 343-365.

[3] E. A. Bender, E. R. Canfield, The asymptotic number of non-negative integer matrices with given row and column sum, *J. Comb. Theory* Ser. A **24** (1978), 296–307.

[4] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *Europ. J. Combinatorics* **1** (1980), 311–316.

[5] R. B. Boppana, J. H. Spencer, A useful elementary correlation inequality, *J. Comb. Theory* A **50** (1989), 305–307.

[6] D. Borkowitz, The name of the game: exploring random permutations, *Mathematics Teacher* **98** (2005) (October)
and its appendix
http://faculty.wheelock.edu/dborkovitz/articles/ngm6.htm

[7] L. Chen, Poisson approximation for independent trials, *Ann. Probab.* **3** (1975), 534–545.

[8] P. Erdős, Graph theory and probability, *Canad. J. Math.* **11** (1959), 34–38.

[9] P. Erdős, J. Kaplansky, The asymptotic number of Latin rectangles, *Amer. J. Math.* **68** (1946), 230–236.

[10] P. Erdős and L. Lovász, Problems and results on 3-chromatic hypergraphs and some related questions, in *Infinite and Finite Sets*, A. Hajnal et. al., Eds., *Colloq. Math. Soc. J. Bolyai* **11**, North Holland, Amsterdam, 1975, 609–627.

[11] P. Erdős and J. H. Spencer, Lopsided Lovász local lemma and latin transversals, *Discrete Appl. Math.* **30** (1991), 151–154.

[12] C. D. Godsil, B. D. McKay, Asymptotic enumeration of Latin rectangles, *J. Comb. Theory* B **48** (1990), 19–44.

[13] C. Y. Ku, Lovász local lemma, `http://www.maths.qmul.ac.uk/~cyk/`

[14] L. Lu and L. A. Székely, Using Lovász Local Lemma in the space of random injections, *Electronic J. Comb.* **14** (2007), R63.

[15] B. D. McKay, Asymptotics for symmetric 0-1 matrices with prescribed row sums, *Ars Combinatoria* **19A** (1985), 15–25.

[16] B. D. McKay and N. C. Wormald, Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$, *Combinatorica* **11** (1991), 369–382.

[17] J. Riordan, Three-line rectangles, *Amer. Math. Monthly* **68** (1946), 230–236.

[18] J. H. Spencer, *Ten Lectures on the Probabilistic Method*, CBMS **52**, SIAM, 1987.

[19] I. Skau, A note on the asymptotic number of Latin rectangles, *Europ. J. Combinatorics* **19** (1998), 617-620.

[20] C. Stein, Asymptotic evaluation of the number of Latin rectangles, *J. Comb. Theory* A **25** (1978), 38–49.

[21] K. Yamamoto, On the asymptotic number of Latin rectangles, *Japan J. Math.* **21** (1951), 113–119.

[22] N. C. Wormald, The asymptotic distribution of short cycles in random regular graphs, *J. Comb. Theory* Series B **31** (1981), 168–182.

[23] N. C. Wormald, Models of random regular graphs. *Surveys in combinatorics,* 1999 (Canterbury), 239–298, London Math. Soc. Lecture Note Ser., **267**, Cambridge Univ. Press, Cambridge, 1999.